<CT> The fragmentation of vision

<CAU> Nick Chater & Ivo Vlaev

<CFN> Nick Chater, Professor of Behavioural Science, Warwick Business School, University of Warwick, Coventry, CV4 7AL, UK. Email: nick.chater@wbs.ac.uk

Ivo Vlaev, Professor of Behavioural Science, Warwick Business School, University of Warwick, Coventry, CV4 7AL, UK. Email: ivo.vlaev@wbs.ac.uk.

<AT>Abstract

<ATX> When looking at a painting, we have the sense of "loading" its contents, in full color and detail, into conscious experience. The physiology and psychology of vision shows that the reality is very different. We "take in" a picture in a series of snippets, as our eyes scan the image. The illusion of being able fully to grasp the whole arises because, almost as soon as we query some aspect of color or detail, our eyes race to find the answer so quickly that it seems to have been available all along. The information is, at it were, at our "visual fingertips." Thus, we spend time looking at a painting not merely to interrogate its meaning, but to explore its basic visual properties.

<TX-N> Why does looking at a painting take so long? Looking at Breughel the Elder's Landscape with the Fall of Icarus (Royal Museums of Fine Arts of Belgium, Brussels), for example, we feel that we are somehow taking in the entire painting within a single visual "gulp." We see the steep hillside, the ploughman blithely tilling his field, the great galleons and placid sea below; the rich red of the ploughman's shirt, the pale green of the sea, the whitish-yellow of the evening sky. All of this, and much more, seems somehow to fuse into a complex, fascinating and harmonious whole, a loaded, in its entirety, into our conscious experience. So why do we continue to look, to examine, to scrutinize, and to ponder? Have we not mentally "hoovered up" Breughel's painting within little more than a glance?

<TX> It is natural to suspect that the purpose of extended looking is about direction our attention and our powers of analysis. Unreflectively, we may feel we "see" the entire painting at once. But the focus of our interest shifts, now to the galleon, now to the ploughman, now to his horse (which we may not previous have noted), then to the shepherd and his staff and, now we come to notice it, a scatter of sheep, and then out to sea. Eventually, we notice a splash, and two flailing legs disappearing into the sea; this is Icarus, as it turns out, who has foolishly flown too close to the sun (Breughel's painting is, admittedly, unusual in presenting its principal subject quite so discreetly). This tragedy occurs in an otherwise peaceful

scene, and seems to pass unnoticed. In Auden puts it in his poem Musée des Beaux Arts [1],

> …the expensive delicate ship that must have seen
> Something amazing, a boy falling out of the sky,
> Had somewhere to get to and sailed calmly on.
> Yet such reflections steal upon us only gradually.

Seeing the painting for the first time, indeed, one might fail to notice Icarus, though central figure in the narrative depicted by the painting, entirely---lost among the clutter of seemingly minor details.

It is easy to imagine that, of course, we "see" all such details merely by looking at the painting---but that we can only pay attention to, and draw implications from, aspects of our conscious experience one by one. This could scarcely be further from the truth. One clue that our subjective sense of visually "grasping" an entire painting must be an illusion is that the sensitivity of our visual system is extremely unevenly distributed across the visual field. The retina has two kinds of light-sensitive receptors: cones, which allow us to pick up color and fine detail, and rods, which are primarily specialized in the detection of change and motion. The cone cells are heavily concentrated in the fovea, a small pit of very densely packed cone cells, which give us high definition and color in a region of roughly five degree of visual angle---with precise detail being picked-up in a small region at its center (the foveola), over a region of just one degree of visual angle [2]. Thus, while standing before Breughel's painting, the sense of simultaneously loading its full color and detail into our visual experience is a fake: we just don't have the retinal 'machinery' to hoover up all this information in a single moment. Indeed, outside the five degrees of visual angle picked up by the fovea we are almost entirely color blind---and our ability to pick out fine detail is extremely more modest.

The limits of our vision are, indeed, much more radical still. A particularly striking illustration comes from an experimental method call gaze-contingent eye tracking. The experimental participants look at the computer screen displaying the image of interest, while their gaze is accurately monitored by an eye tracker. The contents of the computer screen can almost instantaneously be modified, each time a person's eyes jump to a new location in the image (the eye typically moves in discrete saccades, rather than in continuous motion). With this paradigm, it is possible, for example, to display a line of, say English, text where letters are visible only in a 15 character "window" around the current fixation point; everywhere else, words are replaced by blocks of *x*s [3].

If we could load visual impression an entire line of text into our conscious experience, we would have the peculiar impression of seeing a coherent chunk of English hopping around against the background of *x*s. In fact, though, we see nothing unusual. Indeed, reading proceeds entirely normally, and the participant has the impression that they are

looking at an entire sentence, rather than a sequence of partial fragments. Another clue to the astonishing limits of human vision is given by the fact, in normal reading, our eyes hop along the line of text roughly one word at a time (sometimes skipping short, highly predictable, words; and sometimes taking two glimpses at long and unpredictable words). Indeed, roughly speaking, we can only identify one word at a time [4] --- yet, glancing at the text we have a strong impression of simultaneous visual awareness of entire paragraphs. Yet almost all of these paragraph (where we are not directing our foveola) could be replaced by *x*s, or gibberish and we would never now it.

The way brains process scenes, whether photographic or painted, parallels that for reading. It turns out, for example that, using gaze contingent eye tracking as before, images can be doctored such that large, highly visible objects (perhaps one of Breughel's galleons) can blink alternately in and out of existence each time we move our eyes [5]. And we will never notice, until we happen to make an eye movement within, or sufficiently near to, that object [6]. Just as our brain represents just one word at a time, it seems that our brain can only represent one object the time (although an object might be the entire form of the ploughman, the shape of one of his shoes, or a strut of his plough). Indeed, in the rare neuropsychological disorder of simultagnosia, the illusion of visual complexity is undone: people with this condition report seeing only one object at a time. The rest of the visual field is simply blank. It appears that simultagnosia arises when the brain's ability to disengage from the current object, and move rapidly and fluently to the next, is disrupted; and the brain can no longer weave the illusory impression of a full and rich visual scene [7].

The subjective impression that a painting, or our entire surroundings, are loaded in full color and detail into our consciousness has been termed the Grand Illusion [8]. The neuroscientist Donald MacKay [9] brilliantly illustrated how the illusion works, by considering the parallel with touch. Consider, he asked, what it feels like to hold an everyday object, with one's eyes closed. We can sense the continuity and solidity of an object: whether it is rigid (like a phone, a pen or a tennis racket) or flexible (like a notepad or a piece of flex).

MacKay notes how strange it is that we 'feel' the entire object and its movement, even though our fingers are in contact with it at only a few points---and, indeed, as we move turn the object over in our hands, even those 'touch-points' themselves are continually changing. How it is that we 'feel' the entire object, rather than a few 'spots' of disconnected sensation?

MacKay suggested that the felt existence of the object lies not just in the current momentary sensations at the fingers, but in our sense of their *potentialities*. That is, our internal mental model of, say, a ruler or a tennis racket, tells us what we should expect if we start to manipulate it—the anticipated forces, twists, judders, from potential manipulations. The sense of the reality of the whole is confirmed each

time these expectations are confirmed. But, of course, that sense is immediately violated, or at least, transformed, if they are not. Recall, for example, the shock we have when picking up a cup and finding the handle comes away in our hand; fumbling along a solid wall only to across an unexpected doorway; or coming to the bottom of a staircase only to find there is an unexpected extra step, or one fewer than we envisaged.

The essential point is that we 'experience' the whole object or scene, when we are clearly merely touching, or brushing against, tiny fragments of it at any moment. But seeing is no different: our fovea 'touches' a text, scene or painting; it "lands" fleetingly on some particular element of the scene before us (now Icarus's pale legs, now the rigging of the great galleon). And, as with touch, we can actively explore the visual world at will, exploring one element of the scene after another in rapid succession.

Thinking about the purpose of perception, from a moment, this all makes good sense. The purpose of perception is, among other things, to tell us what in the world around us---the layout of words, faces, objects, patterns that surround us. And this world is, of course, defined in precise detail and full color, irrespective of where we happened to be looking at the time, or, for that matter, whether our eyes are open or closed, or whether we are even present at all.

Thus, the brain has the goal of telling us how the world (or a work of art) is---not reporting irrelevant details of where we are looking [10]. So our brain "tells us" that the perceptual world is solid, stable, richly detailed, and highly colored. It does not let us in on the secret that we are viewing the world through a narrow spotlight of precision and color, surrounded by deepening grey fog. If it did so, we would have the sense of the outside, and Breughel's painting, as undergoing remarkable changes as we scan our eyes across it---some items would suddenly snap into focus and color, while others would become vague and grey. But this would, of course, be entirely misleading--- our experience would be suggesting wild flux even as we scan and examine and utterly still painting or scene.

In the light of the sequential nature of vision, it is clear that looking at a painting must inevitably takes a significant amount of time---a glance will pick up no more than one or a few visual fragment. How long do we actually look? Two studies of museum goers, one at the Metropolitan Museum of Art [11] and the other at the Art Institute of Chicago [12] found that typical looking times at well-known paintings are fairly long, but enormously variable. The modal (i.e., most common) looking time was about ten seconds---enough, perhaps, roughly to parse the painting, and to decide that it holds no great interest. On the other hand, viewers will sometimes spend a great deal of time and attention, perhaps up to a few minutes, scrutinizing paintings that particularly attract their attention.

When scanning an individual photograph or painting, our gaze falls very unevenly, typically focusing heavily on people, and especially faces, as our brain actively attempts to reconstruct the scene. Looking at a single face, our gaze hops, primarily, in a triangular region composing eyes and the mouth, which are, of course, particularly indicative of both identity and emotional expression [13]. In a typical viewing episode, though, most parts of most pictures will never directly be inspected at all. This pattern is consistent with the skewed distribution of attention and other resources that is typical in exploration and search problems of all types. So, for example, we scan most news articles only briefly, while reading only few in depth; we skip most web-pages but stop and examine a few in more detail [14]; and we abandon many downloaded pieces of music rapidly, while listening to end only of a few [15]. And considering people in aggregate, of course, the same skewed distribution of "attention" also occurs: a small number of paintings, movies, books, or musical compositions receive a huge amount of attention; the vast majority are lost in obscurity and scarcely attended to by anyone. Indeed, such lop-sided patterns even have a characteristic mathematical form (a so-called "power law"), which is observed for music downloads, books sales, movie receipts and many more [16].

For art forms which are inherently extended in time, such as music, film, or the novel, such careful allocation of our time only to the most interesting examples makes sense. But, as we've seen, viewing a painting is also a temporally extended act and necessarily so. Despite our intuitions to the contrary, rather than taking in an entire painting a single glance, we experience it as a series---perhaps a long series---of glimpses. It seems that not all paintings justify the time and effort required not merely to interrogate its meaning, but to explore its basic visual properties.

What can we conclude? That our sense of having a complete conscious experience of Breughel's painting, or any other, is a fake, conjured out of an extended stream of visual fragments. It is, though, a fake that is created with good reason, to convey correctly the impression of a stable an external world. To return to our initial question: the operation of vision is, in part, why looking at a painting takes so long. Despite our phenomenological intuitions to the contrary, we cannot even see, let along interpret, a painting in a single glance. Breughel is, of course, an illusionist, turning paint into an imitation of a three-dimensional world; and our brain too is an illusionist, creating a subjective sense of a rich, colorful, unchanging world from fleeting snippets of visual experience.

<ACKH>Acknowledgments

<ACK>We would like to thank Martin Kemp, Nicholas Wade, and Charlotte Klonk for valuable suggestions on the manuscript. NC was supported by the ESRC Network for Integrated Behavioural Science [grant number ES/P008976/1].

<REFH>References

<REF>1.
Auden, W. H. (1940). *Another Time*. New York, NY: Random House.

2.
Tschulakow, A. V., Oltrup, T., Bende, T., Schmelzle, S., & Schraermeyer, U. (2018). The anatomy of the foveola reinvestigated. *PeerJ*, *6*, e4482.

3.
McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, *17*(6), 578-586.

4.
Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the EZ Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1-56.

5.
Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception* (*Vancouver Studies in Cognitive Science Vol. 5*) (pp. 89–110). Oxford, UK: Oxford University Press.

6.
Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*(5), 438–443.

7.
Kinsbourne, M. & Warrington, E.K. (1962). A disorder of simultaneous form perception. *Brain, 85*(1), 461–486. Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193-222.

8.
Noë, A., & O'Regan, J. K. (2000). Perception, attention, and the grand illusion. *Psyche*, *6*(15), 6-15.

9.
Mackay, D. M. (1967). Ways of looking at perception. In W. Wathen-Dunn (Ed.). *Models for the Perception of Speech and Visual Form*. Cambridge, MA: MIT Press.

10.
Chater, N. (2018). *The Mind is Flat*. London: Penguin.

11.

Smith, J. K., & Smith, L. F. (2001). Spending time on art. *Empirical Studies of the Arts*, *19*(2), 229-236.

12.

Smith, L. F., Smith, J. K., & Tinio, P. P. (2017). Time spent viewing art and reading labels. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(1), 77-85.

13.
Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press.

14.
Pirolli, P., & Card, S. (1999). Information foraging. *Psychological review*, *106*(4), 643.

15.
Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854-856.

16.
Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, *51*(4), 661-703.